AD 419063

64-5

D1-82-0293

"Also available from the author"

BOEING SCIENTIFIC RESEARCH LABORATORIES

Tchebycheff Approximation and Related Extremal Problems

E. W. Cheney

A. A. Goldstein

Mathematics Research

August 1963

D1-82-0293

TCHEBYCHEFF APPROXIMATION AND RELATED EXTREMAL PROBLEMS

by

E. W. Cheney
University of California, Los Angeles

and

A. A. Goldstein
University of Texas

Visiting Staff Members

Summer 1963

Mathematical Note No. 318

Mathematics Research Laboratory

BOEING SCIENTIFIC RESEARCH LABORATORIES

August 1963

One of the principal maneuvers available for proving existence theorems is to establish that the desired point is where a continuous real-valued function achieves an extremum on a compact set. If the extremum sought is an infimum, then the functional need only be lower semicontinuous. We may take as the definition of this term for a functional φ that each set of the form $\{x : \varphi(x) \leq c\}$ is closed. The crux of an existence proof along these lines would then be the definition of an appropriate topology in which the functional is lower semicontinuous and the set compact. This idea is applied here to several problems involving convex functionals. Applications to Tchebycheff approximation and to control theory are cited. Several theorems dealing with the characterization of extremal points are given as well.

Our terminology follows that of the treatise of Dunford and Schwartz [1] except that we retain the older terms <u>weak</u> and <u>weak* topology</u>. The weak topology in a locally convex space E is defined by saying that a net x_α converges to x if and only if $\langle f,x_\alpha\rangle \to \langle f,x\rangle$ for all $f \in E^*$. Here E^* denotes the space of all continuous linear functionals on E. The weak* topology in E^* is defined by saying that a net f_α converges to f if and only if $\langle f_\alpha,x\rangle \to \langle f,x\rangle$ for all $x \in E$.

<u>§1. Theorem</u>. A continuous convex functional defined on an open convex set in a linear topological space is the supremum of a family of continuous affine functions.

Proof. Let φ be such a functional, G its domain, E the linear space, and R the reals. In $E \oplus R$ consider the upper graph set of φ:

$$K = \{(x,\lambda) : x \in G, \lambda \in R, \lambda \geq \varphi(x)\}.$$

Since φ is upper semicontinuous, K has an interior; since φ is convex, K is convex; since K is lower semicontinuous, K is closed. By [2, page 72, Corollary], K is the intersection of a family of half spaces H_α, say

$$H_\alpha = \{(x,\lambda) : f_\alpha(x) - c_\alpha \lambda \leq d_\alpha\}$$

where α is from some index set A. We note first that $c_\alpha \geq 0$, for in the contrary case, with fixed x and large λ, (x,λ) will lie in K but not in H_α. Let $A^+ = \{\alpha \in A : c_\alpha > 0\}$, and let x be arbitrary in G. Then the inequality $\lambda \geq \varphi(x)$ implies the inequalities $f_\alpha(x) - c_\alpha \lambda \leq d_\alpha$ for all $\alpha \in A^+$. Conversely, if $f_\alpha(x) - c_\alpha \lambda \leq d_\alpha$ for all $\alpha \in A^+$ while $\lambda < \varphi(x)$, then $(x,\lambda) \notin K$ and consequently $f_\beta(x) - c_\beta \lambda > d_\beta$ for some $\beta \in A$. Clearly $\beta \notin A^+$. Thus $c_\beta = 0$. Hence $f_\beta(x) > d_\beta$. But $(x,\varphi(x)) \in K$ and so $f_\beta(x) \leq d_\beta$, a contradiction. We have therefore shown that

$$\lambda \geq \varphi(x) \quad \Leftrightarrow \quad f_\alpha(x) - c_\alpha \lambda \leq d_\alpha \quad (\alpha \in A^+)$$

Hence $\varphi(x) = \sup_{\alpha \in A^+} c_\alpha^{-1}[f_\alpha(x) - d_\alpha]$. ∎

§2. Definition. The c-level set of a functional φ is the set $\{x : \varphi(x) \leq c\}$.

§3. Lemma. A lower semicontinuous functional defined on a locally

convex space and having convex level sets is weakly lower semicontinuous.

<u>Proof</u>. Each set $M_c = \{x : \varphi(x) \leq c\}$ is closed and convex, hence by [1, page 422, theorem 13] weakly closed. Thus φ is weakly lower semicontinuous. ∎

<u>§4. Lemma</u>. A convex functional defined on a linear topological space and having two bounded non-void level sets has all of its level sets bounded.

<u>Proof</u>. Let φ be such a functional. By hypothesis there exists a number a such that the set $S = \{x : \varphi(x) \leq a\}$ is bounded and non-void and contains a point x_o with $\varphi(x_o) < a$. Let $b = a - \varphi(x_o)$. The functional $\Phi(x) = \varphi(x) - \varphi(x_o)$ is convex and satisfies $\Phi(x_o) = 0$. Furthermore $S = \{x : \Phi(x) \leq b\}$. Let M be an arbitrary level set of Φ, say $M = \{x : \Phi(x) \leq c\}$. If $c \leq b$ then $M \subset S$ and M is bounded. If $c > b$ then $0 < \frac{b}{c} < 1$. Hence for arbitrary $x \epsilon M$,

$$\Phi(\frac{b}{c} x + \frac{c-b}{c} x_o) \leq \frac{b}{c} \Phi(x) \leq b$$

Thus the point $\frac{b}{c} x + \frac{c-b}{c} x_o$ lies in S, and x lies in a set of the form $\lambda x_o + \mu S$, with λ and μ independent of x. ∎

<u>Remark</u>. It was pointed out by R. T. Rockafellar that the conclusion requires only <u>one</u> bounded non-void level set in a finite-dimensional setting. An example due to him showing the necessity of two bounded level sets in general is an follows: In ℓ^2 let $\varphi(x) = \Sigma(\frac{x_n}{n})^2$. Then $\{x : \varphi(x) \leq 0\} = \{0\}$, but $\{x : \varphi(x) \leq 1\}$ contains, for each n, the

point $(0,\ldots,0,n,0,\ldots)$ having n for its n^{th} coordinate.

§5. Definition. Given a real-valued function φ defined on a set M, the φ-topology of M has for a subbase all sets of the form $M_c = \{x : \varphi(x) > c\}$.

§6. Theorem. φ achieves its infimum on M if and only if M is compact in the φ-topology.

Proof. Evidently φ is lower semicontinuous in the φ-topology. Hence if M is compact φ will achieve its infimum. For the converse, suppose that φ achieves its infimum μ at x_o. We show M compact. By Alexander's theorem [3, page 139] it suffices to show that any cover by subbase elements has a finite subcover. Suppose then that $M \subset \cup\{M_\lambda : \lambda \epsilon \Lambda\}$ where Λ is a set of reals. We will be finished if there is a $\lambda_o \epsilon \Lambda$ such that $\lambda_o < \mu$, for the $M \subset M_{\lambda_o}$. If no such λ_o exists then $\lambda \geq \mu$ for all $\lambda \epsilon \Lambda$. But then $\varphi(x_o) > \mu$, a contradiction. ∎

§7. Existence Theorem. Let E be a reflexive Banach Space, φ a lower semicontinuous functional on E having bounded convex level sets, and K a closed convex set in E. Then φ achieves its infimum on K.

Proof. By §3, φ is weakly lower semicontinuous. By [1, page 422, theorem 13] K is weakly closed. Thus the set $M = \{x : x \epsilon K, \varphi(x) \leq c\}$ is bounded and weakly closed. By [1, page 425, Corollary 8], M is weakly compact. Hence φ achieves its infimum on K. ∎

§8. Theorem. Let there be given two convex functionals defined on the conjugate of a normal linear space E:

$$\varphi(f) = \sup_{x \in X} \{f(x) - \lambda(x)\}, \quad \psi(f) = \sup_{y \in Y} \{f(y) - \mu(y)\}$$

where X and Y are subsets of E and λ and μ are functionals on X and Y respectively. If for some $c_1 > \inf \varphi$ and $c_2 > \inf \psi$ the c_1-level set of φ and the c_2-level set of ψ have a bounded non-void intersection, then each functional achieves its infimum on each non-void level set of the other.

Proof. A set of the form $\{f : f(x) - \lambda(x) \leq c\}$, with x fixed in E, is closed in the weak* topology. Each level set of φ, being an intersection of such sets, is weak* closed. Similarly for ψ. This proves that φ and ψ are weak* lower semicontinuous. Now suppose that the set of f satisfying $\varphi(f) \leq c_1$ and $\psi(f) \leq c_2$ is non-void and bounded. This set is the same as the set of f for which $\max \{\varphi(f) - c_1, \psi(f) - c_2\} \leq 0$, and the latter is therefore bounded. Since $c_1 > \inf \varphi$ and $c_2 > \inf \psi$, there is an $\epsilon > 0$ such that the set $\max \{\varphi(f) - c_1, \psi(f) - c_2\} \leq -\epsilon$ is also bounded and non-void. By §4 the set of f for which $\max \{\varphi(f) - c_1, \psi(f) - c_2\} \leq c$ is also bounded for any c. If $c_3 \leq c_1 + c$ and $c_4 \leq c_2 + c$ then the set of f for which $\varphi(f) \leq c_3$ and $\psi(f) \leq c_4$ is bounded. This set, then, is weak* closed and bounded, hence weak* compact. Hence φ and ψ achieve their infima there. ∎

Remark. In order that the functional φ achieve its infimum on E^* it is sufficient that one level set of φ be bounded and non-void.

We provide next some examples which illustrate how existence may fail for extremal problems.

§9. Example. In ℓ^p with $p > 1$ define $X = \{\delta^n : n = 1,2,\ldots\}$ where $\delta^n = (0,\ldots,0,1,0,\ldots)$. Define λ on X by putting $\lambda(\delta^n) = n^{-1/q}$, where $q^{-1} + p^{-1} = 1$. If f is a continuous linear functional on ℓ^p then $f(x) = \Sigma\, \alpha_n x_n$ where $\alpha \in \ell^q$. The functional λ does not possess a best Tchebycheff approximation in ℓ^q. Indeed, for the particular sequence

$$\alpha_n = \begin{cases} n^{-1/q} & n < N \\ 0 & n \geq N \end{cases}$$

we get $\varphi(f) = \sup_{x\in X} |f(x) - \lambda(x)| = \sup_n |\alpha_n - n^{-1/q}| = N^{-1/q}$. Thus $\inf_f \varphi(f) = 0$. But to attain $\varphi(f) = 0$ we would have to have $\alpha_n = n^{-1/q}$ for all n, and then $\alpha \notin \ell^q$. In this example, the origin does not lie in the relative interior of the convex closure of $X \cup - X$, so that the hypothesis of §16 below is not satisfied.

§10. Example. In the space $L_1[0,1]$ define the linear functional $g(x) = \int_0^1 tx(t)dt$. The hyperplane $\{x : g(x) = 1\}$ has no point closest to the origin, for in order to minimize $\int|x(t)|dt$ while maintaining $g(x) = 1$, the total area under the x-curve would have to be piled up at the point 1. In §14 below a simple sufficient condition is given in order that a linear manifold shall contain a point closest to the origin.

§11. Example. Consider a Markoff system in $C[0,1]$ which is not fundamental. For example $\{x^{2^n}\}$. Let M denote the closed linear manifold spanned by these functions. No point outside M has a closest point in M, for by the Tchebycheff alternation theorem the error curve for such an approximation would have to alternate an infinite number of times between $+\epsilon$ and $-\epsilon$, which is not possible.

§12. Example. In ℓ^2 let X denote the Hilbert "cube," i.e., the set of points x such that $|x_n| \leq 1/n$ for all n. On X define λ by $\lambda(x) = \sum_{n=1}^{\infty} x_n/\sqrt{n}$. The series is absolutely convergent because the series $\Sigma\, n^{-3/2}$ converges. Let $\varphi(f) = \sup_{x \in X} |f(x) - \lambda(x)|$. By taking f to be a sequence

$$a_n = \begin{cases} 1/\sqrt{n} & n \leq N \\ 0 & n > N \end{cases}$$

We get $\varphi(f) = \sup_{|x_n| \leq \frac{1}{n}} \left| \sum_{n>N} \frac{x_n}{\sqrt{n}} \right| \leq \sum_{n>N} n^{-3/2}$. Thus $\inf_f \varphi(f) = 0$. But it is not possible to achieve $\varphi(f) = 0$. Indeed, if $f = (a_1, \ldots)$ then clearly $a_n \neq \frac{1}{\sqrt{n}}$ for some n. Take then $x = \frac{1}{n}\delta_n$ to show $|f(x) - \lambda(x)| = \frac{1}{n}|a_n - \frac{1}{\sqrt{n}}| > 0$. As in the example of §9, the hypothesis of §16 is violated.

§13. Example. In any Banach space E of infinite dimension, let $\{x_n\}$ be a linearly independent sequence of vectors with $\|x_n\| \downarrow 0$. Put $X = \{x_n\} \cup \{0\}$. Then X is compact. Define $\lambda(x_n) = \sqrt{\|x_n\|}$ and $\lambda(0) = 0$. Then λ possesses no best Tchebycheff approximation from E^*. Indeed, for fixed N, if we put $f(x_n) = \lambda(x_n)$ for $n < N$ and $f(x_n) = 0$ for $n \geq N$ then f may be extended linearly to an element of E^*. But $\varphi(f) = \sup_{x \in X} |f(x) - \lambda(x)| = \|x_N\|$. Hence $\inf_{f \in E^*} \varphi(f) = 0$. But

if $\varphi(f_o) = 0$ for some f_o then $f_o = \lambda$ on X, and then f_o is unbounded since $f_o(x_n/\|x_n\|) = \|x_n\|^{-1/2} \to \infty$. As in the example of §9 the hypothesis of §16 is violated.

<u>§14. Theorem</u>. Let E be a normed linear space and M a weak* closed subset of E^*. Then M contains a point of minimum norm.

<u>Proof</u>. The sets $\{f : \|f\| \leq c\}$ are weak* compact by [1, page 424, theorem 2]. Hence the norm in E^* is weak* lower semicontinuous. Since M is weak* closed, each set $\{f \in M : \|f\| \leq c\}$ is weak* compact, and hence $\|f\|$ achieves its infimum. $\blacksquare$

<u>§15. Example</u>. As an illustration of an extremum being achieved we consider a problem of automatic control. There is given a system of linear differential equations

$$(1) \quad \dot{x}_i(t) = \sum_{j=1}^{n} A_{ij}(t)\, x_j(t) + \sum_{j=1}^{n} B_{ij}(t)\, u_j(t) + c_i(t) \qquad (i = 1,\dots,n)$$

or

$$\dot{x} = Ax + Bu + C$$

in which A and B are $n \times n$ matrices depending continuously on $t \in [0,1]$, and $x(t)$, $u(t)$, and $c(t)$ are n-vectors. It is assumed that $c(t)$ is continuous. The control problem consists in finding a control $u(t)$ such that the differential equations above, with prescribed initial and terminal values $x(0)$ and $x(1)$, will have a solution $x(t)$, and such that under this restriction the integral

$$(2) \qquad \int_0^1 [u_1^2(s) + \cdots + u_n^2(s)]^{1/2} ds$$

will be minimized. Since we wish to permit the u_i to be generalized functions, it is more convenient to rephrase the problem after first passing to the explicit solution of the initial value problem. To this end, let $\Xi(t)$ be the matrix of fundamental solutions of (1) such that $\Xi(0)$ is the identity matrix. Then the initial value problem has the solution

$$x_i(t) = \sum_{j=1}^{n} \Xi_{ij}(t)\{x_j(0) + \int_0^t \sum_{k=1}^{n} \Xi_{jk}^{-1}(s)[B_{jk}(s)\, u_k(s) + c_k(s)]ds\}$$

or

$$x(t) = \Xi(t)\{x(0) + \int_0^t \Xi^{-1}(s)[B(s)\, u(s) + c(s)]ds\}$$

Let $u_j(s)ds = dv_j(s)$. The condition that the prescribed terminal values $x_j(1)$ be achieved is that

$$(3) \qquad \int_0^1 \Xi^{-1}(s)B(s)dv(s) = \alpha$$

where α is a numerical n-tuple given by

$$\alpha = \Xi^{-1}(1)x(1) - \int_0^1 \Xi^{-1}(s)\, C(s)ds.$$

Now let $\mathcal{C}$ denote the n-fold direct sum of $C[0,1]$. An element of $\mathcal{C}^*$ may be identified with an n-tuple $(v_1,\ldots,v_n)$ where each v_i is a function of bounded variation on $[0,1]$. For $v = (v_1,\ldots,v_n)\in\mathcal{C}^*$ and $y = (y_1,\ldots,y_n)\in\mathcal{C}$ we set

$$\langle v,y\rangle = \sum_{j=1}^{n} \int_0^1 y_j(s)dv_j(s)$$

Now for $i = 1,\ldots,n$ let Y_i be a point in $\mathcal{C}$ with components

$$Y_{ij} = \sum_{k=1}^{n} \Xi^{-1}_{ik}(s)B_{kj}(s).$$

Then equation (3) takes the form

$$\langle v, Y_i \rangle = \alpha_i \qquad (i = 1,\ldots,n)$$

From this we see that v is constrained to lie on the intersection of n hyperplanes in $\mathcal{C}^*$, whose gradients are from $\mathcal{C}$. The set of all such v is therefore a weak* closed linear manifold.

Now let $\mathcal{C}$ be normed as follows: $\| x \| = \sup\limits_{0\le t\le 1} \sqrt{\Sigma\, x_i^2(t)}$. This norm on $\mathcal{C}$ induces a norm on $\mathcal{C}^*$, and we shall show that for a linear functional v of the form

$$\langle v, x \rangle = \sum_i \int_0^1 x_i(s)\, u_i(s)ds$$

the norm $\| v \|$ is precisely the integral (2).

In fact, by the Cauchy-Schwartz inequality for n-tuples, we have

$$\| v \| = \sup_{x\in\mathcal{C}, \| x \|=1} \langle v, x \rangle$$

$$= \sup_{\| x \|=1} \int_0^1 \sum_i x_i(s)u_i(s)ds$$

$$\le \sup_{\| x \|=1} \int_0^1 \sqrt{\Sigma\, x_i^2(s)}\ \sqrt{\Sigma\, u_i^2(s)}\ ds$$

$$\le \int_0^1 \sqrt{\Sigma\, u_i^2(s)}\ ds$$

On the other hand, we can take as a particular choice for x the function of norm 1 whose components are

$$x_i(s) = u_i(s)/\sqrt{\Sigma\, u_j^2(s)}$$

Then $\|v\| \geq \langle v,x\rangle = \int_0^1 \sqrt{\Sigma\, u_i^2(s)}\, ds.$

If the scope of the problem is enlarged so that the controls are differentials of functions of bounded variation then the existence of optimal controls is guaranteed by the simple theorem of §14.

In the next theorems we are concerned with an abstract problem of Tchebycheff approximation. Given a set X in a linear topological space E and a bounded functional λ on X we seek an element $f \epsilon E^*$ for which the expression

$$\varphi(f) = \sup_{x \epsilon X} |f(x) - \lambda(x)|$$

is an absolute minimum. If such an f exists, it is termed a "best Tchebycheff approximation to λ." This problem encompasses the familiar problems of approximating functions by linear families of other functions. This will be mentioned again later. We begin with an existence theorem. The notation $[X]$ denotes the linear subspace spanned by a set X. The convex hull of X, denoted by $\mathfrak{H}(X)$ is the set of all finite linear combinations $\Sigma\, \theta_i x_i$ where $x_i \epsilon X$, $\theta_i \geq 0$, $\Sigma\, \theta_i = 1$. The convex closure of X is the closure of the convex hull, and is denoted by $\bar{\mathfrak{H}}(X)$. We mean by the relative interior of X its interior relative to the least linear manifold containing X.

§16. Existence Theorem for Tchebycheff Approximation. Let E be a Banach space and X a subset of E. In order that each bounded function on X possess a best Tchebycheff approximation in E^* it is sufficient that, in the relative topology of $[X]$, 0 be interior to the convex closure of $X \cup -X$.

Proof. The space $[X]$ is a normed linear space (possibly incomplete) and $[X]^*$ is a Banach space. We begin by showing that the functional φ defined on $[X]^*$ by the equation $\varphi(f) = \sup_{x \in X} |f(x) - \lambda(x)|$ achieves its infimum. By the remark of §8, it suffices to prove that each level set of φ is bounded. Suppose $\varphi(f) \leq c$. Then $|f(x) - \lambda(x)| \leq c$ for all $x \in X$. Thus $|f(x)| \leq b$ for an appropriate b. This inequality is valid also for $x \in -X$. By linearity the inequality remains true for $x \in \mathfrak{S}(X \cup -X)$, and by continuity for $x \in \bar{\mathfrak{S}}(X \cup -X)$. Thus f is bounded by b on a sphere, and this makes f bounded by a constant independent of f. Thus φ achieves its infimum on $[X]^*$, say at f_o. By the Hahn-Banach theorem, f_o may be extended to an element of E^* without changing $\varphi(f_o)$. On the other hand, any element of E^* becomes an element of $[X]^*$ when restricted to X. Hence f_o (extended) solves the minimization problem. ∎

§17. Characterization Theorem. Let X be a bounded subset of a locally convex space E, and λ a functional on X which is bounded below. In order that a point $f_o \in E^*$ minimize the expression $\varphi(f) = \sup_{x \in X} \{f(x) - \lambda(x)\}$ it is necessary and sufficient

that $0 \in \bigcap_{\epsilon>0} \overline{\mathfrak{S}}X_\epsilon$, where $X_\epsilon = \{x \in X : f_0(x) - \lambda(x) \geq \varphi(f_0) - \epsilon\}$.

Proof. If $0 \notin \overline{\mathfrak{S}}(X_\epsilon)$ then by the separation theorem, there exists an $h \in E^*$ such that $\inf_{x \in \overline{\mathfrak{S}}(X_\epsilon)} h(x) = \delta > 0$. Let c be a positive number such that $-ch(x) < \frac{1}{2}\epsilon$ for all $x \in X$. Then $\varphi(f_0 - ch) < \varphi(f_0)$. Indeed, for $x \in X_\epsilon$ we have $f_0(x) - ch(x) - \lambda(x) \leq \varphi(f_0) - c\delta$, and for $x \in X \sim X_\epsilon$ we have $f_0(x) - ch(x) - \lambda(x) < \varphi(f_0) - \epsilon - ch(x) < \varphi(f_0) - \epsilon/2$.

For the converse, suppose that f_0 does not minimize φ. Then there is an $h \in E^*$ and an $\epsilon > 0$ such that $\varphi(f_0 - h) + 2\epsilon < \varphi(f_0)$. Let y be a point of X_ϵ. Then $\varphi(f_0) \leq f_0(y) - \lambda(x) + \epsilon$, whence $f_0(y) - h(y) - \lambda(y) \leq \varphi(f_0 - h) < \varphi(f_0) - 2\epsilon \leq f_0(y) - \lambda(x) - \epsilon$. Thus $-h(y) < -\epsilon$, and 0 cannot lie in the convex closure of X_ϵ. $\blacksquare$

The analogous theorem for Tchebycheff approximation, in the real or complex case may be stated as follows:

§18. Theorem. Let X be a bounded subset of a locally convex space E over the real or complex field, and λ a bounded functional on X. A necessary and sufficient condition that an element $f \in E^*$ shall minimize the expression $\varphi(f) = \sup_{x \in X} |f(x) - \lambda(x)|$ is that $0 \in \bigcap_{\epsilon>0} \overline{\mathfrak{S}}X_\epsilon$, where $X_\epsilon = \{\overline{r(x)} \cdot (x) : |r(x)| \geq \varphi(f) - \epsilon\}$ and $r = f - \lambda$.

§19. Lemma. Let $\{A_n\}$ be a decreasing sequence of compact sets in a locally convex space. Then $\overline{\mathfrak{S}}\bigcap A_n = \bigcap \overline{\mathfrak{S}}A_n$.

Proof. Put $A = \bigcap A_n$. Since $\bigcap \overline{\mathfrak{S}}A_n$ is a closed convex set containing A,

it contains $\bar{\mathfrak{H}}A$ also.

For the reverse inclusion, note first that $\bar{\mathfrak{H}}A$ is the intersection of all closed half spaces containing A. It will therefore suffice to prove that if a half space $\{x : f(x) \leq c\}$ contains A then it contains $\cap \bar{\mathfrak{H}}A_n$. Given ϵ, there corresponds an m such that $A_m \subset \{x : f(x) \leq c + \epsilon\}$. Indeed in the contrary case, we could find a sequence x_n with the property $x_n \epsilon A_n \cap \{x : f(x) \geq c + \epsilon\}$, and any cluster point x of this sequence would lie in A but not in $\{x : f(x) \leq c\}$. Thus the half space $\{x : f(x) \leq c + \epsilon\}$ contains A_m, hence $\bar{\mathfrak{H}}A_m$, and finally $\cap \bar{\mathfrak{H}} A_n$. Since this is true for every ϵ, the proof is complete. This argument by R. T. Rockafellar replaces our earlier and clumsy proof. ∎

<u>§20. Theorem</u>. Let X be a non-void compact set in a (possibly complex) locally convex space E, and λ a continuous functional on X. In order that an element $f \epsilon E^*$ shall minimize the expression $\varphi(f) = \sup_{x \epsilon X} |f(x) - \lambda(x)|$ it is necessary and sufficient that 0 lie in the convex closure of the set $\{\overline{r(x)} \cdot x : x \epsilon X, |r(x)| = \varphi(f)\}$, where $r = f - \lambda$.

<u>Proof 1</u>. By §18, the necessary and sufficient condition on f is that $0 \epsilon \bigcap_{\epsilon > 0} \bar{\mathfrak{H}} X_\epsilon$ where $X_\epsilon = \{\overline{r(x)} \cdot x : |r(x)| + \epsilon \geq \varphi(f)\}$. Since λ is continuous, X_ϵ is compact. By the lemma, $\cap \bar{\mathfrak{H}} X_\epsilon = \bar{\mathfrak{H}} \cap X_\epsilon$. But $\cap X_\epsilon = X_0$. ∎

<u>Proof 2</u>. A direct proof using the separation theorem is also possible. Assume first that $0 \not\epsilon \bar{\mathfrak{H}} X_0$. Then $\varphi(f) > 0$. By the separation theorem, there exists $h \epsilon E^*$ such that $\alpha \equiv \inf_{x \epsilon \bar{\mathfrak{H}}(X_0)} \mathcal{R}[h(x)] > 0$. Let $Z = \{x \epsilon X :$

$|r(x)| > \frac{1}{2}\varphi(f)$ and $\mathcal{R}[\overline{r(x)}\ h(x)] > \frac{1}{2}\alpha\}$. Since Z is open in X, the set $Y = X \sim Z$ is compact. Consequently the supremum of $|r(y)|$ is attained on Y, say at y_o. If $p \equiv |r(y_o)| = \varphi(f)$ then $r(y_o)\cdot y_o \epsilon X_o$ and $y_o \epsilon Z$, contradicting the choice of y_o. Hence $p < \varphi(f)$. Now let $\beta = 1 + \max\limits_{x\epsilon X}|h(x)|$ and let $0 < \Theta < \frac{1}{2}\alpha/\beta^2$. If $x\epsilon Z$ then $|r(x) - \Theta h(x)|^2 = |r(x)|^2 + \Theta^2|h(x)|^2 - 2\Theta\,\mathcal{R}[h(x)\ \overline{r(x)}] \leq |r(x)|^2 + \Theta^2\beta^2 - \Theta\alpha < [\varphi(f)]^2 - \frac{1}{2}\alpha\Theta$. If $0 < \Theta < \beta^{-1}[\varphi(f) - p]$ then for $x\epsilon Y$, $|r(x) - \Theta h(x)| \leq |r(x)| + \Theta|h(x)| \leq p + \Theta\beta < \varphi(f)$. Consequently for some Θ, $\varphi(f - \Theta h) < \varphi(f)$.

For the converse, assume that $\varphi(f - h) < \varphi(f)$. Then $|(f - h)(x) - \lambda(x)| < |f(x) - \lambda(x)|$ for $x\epsilon S = \{x\epsilon X : |r(x)| = \varphi(f)\}$. Hence $|(f - h)(x) - \lambda(x)|^2 = |r(x)|^2 + |h(x)|^2 - 2\mathcal{R}[h(x).\overline{r(x)}] < |r(x)|^2$, and $\mathcal{R}[h(x)\cdot\overline{r(x)}] > \frac{1}{2}|h(x)|^2$. Since $h(x)$ does not vanish on S and since S is compact, $\mathcal{R}[h(x).\overline{r(x)}]$ is bounded away from 0 on S, and consequently, $0 \notin \mathfrak{S} X_o$. This Theorem was given in the real case in a research announcement, Bulletin of the Am. Math. Soc. [68, (1962), 449-450]. ▮

§21. Definition. The cone of a set Y consists of the origin together with all the positive multiples of points in Y. Thus in particular the cone of the void set is $\{0\}$. The cone of Y is denoted by $\mathcal{C}(Y)$.

§22. Lemma. In a linear topological space the cone of a closed and bounded set not containing 0 is closed.

<u>Proof</u>. Let Y be such a set and E the space. If $\mathcal{C}(Y)$ is not closed then there is a point $u \in E \sim \mathcal{C}(Y)$ such that every neighborhood of u meets $\mathcal{C}(Y)$. Then to each neighborhood α of 0 there will correspond $u_\alpha \in (u+\alpha) \cap \mathcal{C}(Y)$. We may write $u_\alpha = \lambda_\alpha y_\alpha$ where $\lambda_\alpha \geq 0$ and $y_\alpha \in Y$. By passing to a subnet we may suppose that λ_α converges, possibly to $+\infty$. In the latter case, $y_\alpha \to 0$ contradicting the hypothesis $0 \notin Y$. If $\lambda_\alpha \to 0$ then $u_\alpha \to 0$ because Y is bounded. Since $u \neq 0$ this possibility is ruled out also. Consequently, $\lambda_\alpha \to \lambda$ with $0 < \lambda < \infty$. Hence $y_\alpha \to u/\lambda$. Since Y is closed $u/\lambda \in Y$, contradicting $u \notin \mathcal{C}(Y)$. ∎

<u>§23. Lemma</u>. Let Y be a closed and bounded convex set having a closed cone in a locally convex space, and let X be a compact convex set disjoint from the cone of Y. Then there exists a continuous linear functional f such that

$$\sup_{x \in X} f(x) < 0 < \inf_{y \in Y} f(y)$$

<u>Proof</u>. By the ordinary separation theorem, there exists a continuous linear functional h such that $\inf_{y \in Y} h(y) > 0$. Since $\mathcal{C}(Y)$ is closed, there exists a continuous linear functional g such that

$$\sup_{x \in X} g(x) < 0 \leq \inf_{y \in \mathcal{C}(Y)} g(y)$$

Now let $0 < \lambda < \inf_{\xi \in X} [-g(\xi)] \div \sup_{x \in X} 1 + |h(x)|$, and define $f = g + \lambda h$. We have then for $x \in X$, $f(x) = g(x) + \lambda h(x) < g(x) + \lambda[1 + |h(x)|] \leq g(x) + \inf_{\xi \in X} [-g(\xi)] \leq 0$. Since X is compact and f is continuous, $\sup_{x \in X} f(x) < 0$.

On the other hand, $\inf_{y\in Y} f(y) = \inf_{y\in Y} [g(y) + \lambda h(y)] \geq \inf_{y\in Y} \lambda h(y) > 0.$ ∎

§24. Existence Theorem. Let X and Y be subsets of a normed linear space E. Let λ and μ be bounded functionals on X and Y respectively. If, in the relative topology of $[X\cup Y]$, 0 is interior to the convex closure of $X\cup Y$ then the functional $\varphi(f) = \sup_{x\in X} [f(x) - \lambda(x)]$ achieves its infimum on the set $K = \{f\epsilon E^* : \sup_{y\in Y} [f(y) - \mu(y)] \leq 0\}$.

Proof. We proceed as in §16. For any constant c the set S of $f\epsilon[X\cup Y]^*$ for which $\sup_{y\in Y} [f(y) - \mu(y)] \leq 0$ and $\sup_{x\in X} [f(x) - \lambda(x)] \leq c$ is bounded. Indeed, if $f\epsilon S$ then because of the boundedness of λ and μ, $f(z) \leq k$ for some k and all $z\epsilon X\cup Y$. This inequality remains true for $z\epsilon \bar{\mathfrak{H}}(X\cup Y)$ and hence on a sphere about 0. Thus S is bounded. By familiar arguments, S is compact in the weak* topology of $[X\cup Y]^*$, and φ is weak* lower semicontinuous. An element of $[X\cup Y]^*$ which minimizes φ may be extended by the Hahn-Banach theorem to E^*. ∎

Remark. It is only necessary to assume that for some $X_o \subset X$ and $Y_o \subset Y$, μ is bounded on Y_o, λ is bounded on X_o, and 0 is in the interior of $\bar{\mathfrak{H}}(X_o \cup Y_o)$ relative to $[X\cup Y]$.

§25. Characterization Theorem for Convex Programming. Let E be a locally convex space, X a bounded subset of E, and Y a compact subset of E with $0 \notin \bar{\mathfrak{H}}(Y)$. Let λ be a functional on X which is bounded below. Let μ be a lower semicontinuous functional on Y. Put $K = \{f\epsilon E^* : f(y) \leq \mu(y)$ for all $y\epsilon Y\}$

and $\varphi(f) = \sup\{f(x) - \lambda(x) : x \epsilon X\}$. In order that a point f_0 of K minimize φ on K it is necessary and sufficient that for each $\epsilon > 0$, $\mathcal{C}\bar{\mathfrak{H}}(-Y_0) \cap \bar{\mathfrak{H}}(X_\epsilon) \neq \emptyset$ where $Y_0 = \{y \epsilon Y : f_0(y) = \mu(y)\}$ and $X_\epsilon = \{x \epsilon X : f_0(x) - \lambda(x) \geq \varphi(f_0) - \epsilon\}$.

<u>Proof</u>. (Sufficiency) Suppose that f_0 does not minimize φ on K. Then for some $h \epsilon E^*$, $f_0 - h \epsilon K$ and $\varphi(f_0 - h) < \varphi(f_0)$. Select $\epsilon > 0$ such that $\varphi(f_0 - h) < \varphi(f_0) - 2\epsilon$. For $y \epsilon Y_0$ we have $(f_0 - h)(y) \leq \mu(y)$, whence $h(y) \geq 0$. This inequality persists if y ranges over $\mathcal{C}\bar{\mathfrak{H}}(Y_0)$. For $x \epsilon X_\epsilon$ we have $(f_0 - h)(x) - \lambda(x) \leq \varphi(f_0 - h) < \varphi(f_0) - 2\epsilon < f_0(x) - \lambda(x) - \epsilon$, whence $h(x) \geq \epsilon$. This inequality persists if x ranges over $\bar{\mathfrak{H}}(X_\epsilon)$. Thus the sets $\mathcal{C}\bar{\mathfrak{H}}(-Y_0)$ and $\bar{\mathfrak{H}}(X_\epsilon)$ are disjoint.

(Necessity) Suppose that the sets $\mathcal{C}\bar{\mathfrak{H}}(-Y_0)$ and $\bar{\mathfrak{H}}(X_\epsilon)$ are disjoint for some $\epsilon > 0$. By §23, there exists an $h \epsilon E^*$ and constants α and β such that

$$h(y) \leq 2\alpha < 0 < \beta \leq h(x)$$

for all $y \epsilon \bar{\mathfrak{H}}(-Y_0 \cup \{0\})$ and all $x \epsilon \bar{\mathfrak{H}}(X_\epsilon)$. Since $X \cup Y$ is bounded there exists a number k such that $|h(x)| \leq k$ for all $x \epsilon X \cup Y$. Now let $Y^* = \{y \epsilon Y : h(y) > -\alpha\}$. Y^* is then an open set in Y containing Y_0. Since $Y \sim Y^*$ is compact and contains no points of Y_0, the number

$$\gamma = \max_{y \epsilon Y \sim Y^*} \{f_0(y) - \mu(y)\}$$

is negative. Select a positive number t such that $2tk < \max\{\epsilon, -\gamma\}$. We are going to show that $f_0 - th$ is a point of K yielding a lower

value of φ than $\varphi(f_o)$. For $y \in Y^*$ we have $(f_o - th)(y) - \mu(y) \leq - th(y) \leq 0$. For $y \in Y \sim Y^*$ we have $(f_o - th)(y) - \mu(y) \leq \gamma + tk < \frac{1}{2}\gamma \leq 0$. Thus $f_o - th \in K$. For $x \in X_\epsilon$ we have $(f_o - th)(x) - \lambda(x) \leq \varphi(f_o) - th(x) < \varphi(f_o) - t\beta$. For $x \in X \sim X_\epsilon$ we have $(f_o - th)(x) - \lambda(x) \leq \varphi(f_o) - \epsilon - th(x) \leq \varphi(f_o) - \epsilon + tk < \varphi(f_o) - \frac{1}{2}\epsilon$. Hence $\varphi(f_o - th) < \varphi(f_o)$. ∎

Remark. The hypothesis $0 \notin \bar{\mathfrak{H}}(Y)$ may be replaced by the hypothesis that $\mathcal{C}\bar{\mathfrak{H}}(Y)$ is closed.

§26. Corollary. If, in the above theorem, E is complete, X is compact, and λ is lower semicontinuous, then the necessary and sufficient condition on f_o is that $\mathcal{C}\bar{\mathfrak{H}}(-Y_o) \cap \bar{\mathfrak{H}}(X_o)$ shall be non-void.

Proof. From the theorem, the necessary and sufficient condition on f_o is that for each $\epsilon > 0$ there exist a point v_ϵ in $\mathcal{C}\bar{\mathfrak{H}}(Y_o) \cap \bar{\mathfrak{H}}(X_\epsilon)$. Since $\bar{\mathfrak{H}}(X_\epsilon)$ is compact [2, page 81, corollary], the set $\{v_\epsilon : \epsilon > 0\}$ has a point of accumulation, v. Clearly $v \in \bigcap_{\epsilon > 0} \bar{\mathfrak{H}} X_\epsilon$. But by §19, $\bigcap \bar{\mathfrak{H}} X_\epsilon = \bar{\mathfrak{H}} \bigcap X_\epsilon = \bar{\mathfrak{H}} X_o$. ∎

§27. Example. In the linear space C[0,1] we introduce the norm $\| x \| = \sup_t (t+1)|x(t)|$, which is topologically equivalent to the usual norm. Consider the problem of minimizing $\| f \|$ in the conjugate space under the constraint $f(I) = 1$ where I denotes the function identically 1 on [0,1]. The theorem of §25 may be applied to prove that a solution is given by $f_o(x) = x(1)$. To this end, let X denote the unit sphere in C and let $Y = \{\pm I\}$. We set $\mu(\pm I) = \pm 1$

and $\lambda(x) = 0$. Then $K = \{f : f(I) = 1\}$ and $\varphi(f) = \|f\|$. The set Y_o is $\{\pm I\}$ and the set X_ϵ is $\{x \in C : x(1) \geq 1 - \epsilon\}$. Since $I \in Y_o \cap Y_\epsilon$, the condition of the theorem is met, and f_o is a solution.

<u>§28. Theorem</u>. Let E_1 and E_2 be Banach spaces. Let X be a subset of E_1 such that $0 \in \operatorname{int} \mathfrak{C}(X \cup -X)$. Let Λ be a bounded map of X into E_2^*. Then Λ has a best Tchebycheff approximation by an element of $B(E_1, E_2^*)$.

<u>Proof</u>. Define the τ-topology in $B(E_1, E_2^*)$ by saying that $L_\alpha \to L$ iff $\langle y, L_\alpha x\rangle \to \langle y, Lx\rangle$ for all $x \in E_1$ and all $y \in E_2$. We are going to show that the function $\varphi(L) = \sup\limits_{x \in X} \|Lx - \Lambda x\|$ achieves its infimum in $B(E_1, E_2^*)$. It will suffice to show that φ is τ-lower semicontinuous and that each set $M_c = \{L \in B(E_1, E_2^*) : \varphi(L) \leq c\}$ is τ-compact. That M_c is τ-closed is quite obvious, and thus φ is τ-lower semicontinuous. Now M_c is norm bounded because if $L \in M_c$ then $\|Lx - \Lambda x\| \leq c$ for all $x \in X$ whence $\|Lx\| \leq c'$ for all $x \in X$. This inequality remains true for $x \in \mathfrak{C}(X \cup -X)$ and thus in a sphere about 0. Thus $\|L\|$ is bounded by a number depending only on c. Since M_c is norm bounded and τ-closed, we may apply Corollary 3 of [2, chapter 4, page 65] to conclude that M_c is τ-compact. $\blacksquare$

<u>§29. Lemma</u>. Let E and F denote Banach spaces. Let X be compact in E and Y compact in F. In order that there exist a bounded linear transformation $A : F \to E^*$ such that $\langle x, Ay\rangle > 0$ for all $x \in X$ and $y \in Y$, it is necessary and sufficient that the

convex closures of X and Y be disjoint from 0.

<u>Proof</u>. If $\langle x,Ay\rangle > 0$ for all $x \epsilon X$ and $y \epsilon Y$ then by compactness, $x,Ay\rangle \geq \epsilon$ for some $\epsilon > 0$. This remains true for $x \epsilon \bar{\mathfrak{H}} X$ and $y \epsilon \bar{\mathfrak{H}} Y$. Consequently $0 \notin \bar{\mathfrak{H}} X$ and $0 \notin \bar{\mathfrak{H}} Y$.

Conversely, suppose $0 \notin \bar{\mathfrak{H}} X$ and $0 \notin \bar{\mathfrak{H}} Y$. By the separation theorem, there exists $u \epsilon E^*$ and $v \epsilon F^*$ such that $\langle x,u\rangle > 0$ for all $x \epsilon X$ and $\langle y,v\rangle > 0$ for all $y \epsilon Y$. Define $A : F \to E^*$ by putting $\langle x,Ay\rangle = \langle u,x\rangle\langle v,y\rangle$ for arbitrary $x \epsilon E$ and $y \epsilon F$. Clearly A is a bounded linear transformation satisfying $\langle x,Ay\rangle > 0$ for $x \epsilon X$ and $y \epsilon Y$. $\blacksquare$

<u>§30. Theorem</u>. Let X be a compact set in a Banach space E and Φ a continuous map of E into Hilbert space. In order that a bounded linear operator A from E into the Hilbert space shall minimize the expression $\Delta(A) = \sup_{x \epsilon X} \| Ax - \varphi x \|$ it is necessary and sufficient that 0 lie in the convex closure of $X_o \equiv \{x \epsilon X : \| Ax - \varphi x \| = \Delta(A)\}$ or of $\{Ax - \varphi x : x \epsilon X_o\}$.

<u>Proof</u>. Put $Rx = Ax - \varphi x$. If $\Delta(A)$ is not a minimum then for some linear operator B, $\Delta(A-B) < \Delta(A)$. Thus for $x \epsilon X_o$, $\| Ax - \varphi x - Bx \|^2 < \| Ax - \varphi x \|^2$ whence $-2(Rx,Bx) + \| Bx \|^2 < 0$ and $(Rx,Bx) > 0$. By the lemma, 0 is disjoint from the convex closures of the point sets $\{Rx : x \epsilon X_o\}$ and X_o.

For the converse, suppose that 0 is disjoint from the convex closures of the sets $\{Rx : x \epsilon X_o\}$ and X_o. By the lemma, there exists a linear operator B such that $(Rx,Bx) > 0$ for all $x \epsilon X_o$. Since X_o is compact, the number $\epsilon_1 = \min_{x \epsilon X_o} (Rx,Bx)$ is positive. Put $\epsilon_2 = \max_{x \epsilon X} \| Bx \|$.

Define $X_1 = \{x\epsilon X : (Bx,Rx) > \frac{1}{2}\epsilon_1, \ \|Rx\| > \frac{1}{2}\Delta(A)\}$. X_1 is open in X and contains X_o. Consequently $X \sim X_1$ is a compact set containing no points of X_o. Consequently $\|Rx\|$ has a maximum, e, on $X \sim X_1$ such that $e < \Delta(A)$. Let $0 < \lambda < \epsilon_1/\epsilon_2^2$ and $0 < \lambda < \dfrac{\Delta(A) - e}{2\epsilon_2}$. We are going to prove $\Delta(A - \lambda B) < \Delta(A)$. For $x\epsilon X_1$ we have $\|(A - \lambda B)x - \varphi x\|^2 = \|Rx\|^2 - 2\lambda(Rx,Bx) + \lambda^2\|Bx\|^2 < \|Rx\|^2 - \lambda\epsilon_1 + \lambda^2\epsilon_2^2 < \|Rx\|^2 + \lambda\epsilon_2^2(\lambda - \frac{\epsilon_1}{\epsilon_2^2})$. For $x\epsilon X{\sim}X_1$ we have $\|(A - \lambda B)x - \varphi x\| < \|Rx\| + \lambda\|Bx\| < e + \lambda\epsilon_2 < e + \frac{1}{2}[\Delta(A) - e] < \Delta(A)$. ∎

In the next theorem we consider the problem of non-linear Tchebycheff approximation in an abstract setting. Let E be a reflexive Banach space over the real field, and let T be a compactum. Let f be a map of $E \times T$ to the field, continuous in x and t separately. We assume that f has a Frechet derivative f' with respect to the first variable x, that $\{f'(x,t) : t\epsilon T\}$ is equicontinuous and that f' is continuous in t. The points $x\epsilon E$ for which the expression

$$\Delta(x) = \sup_{t\epsilon T} |f(x,t)|$$

is (locally) a minimum are partially characterized by the following theorem. Set $T_x = \{t\epsilon T : |f(x,t)| = \Delta(x)\}$.

<u>§31. Theorem</u>. If a point $x\epsilon E$ is a local minimum point of Δ then the origin of E^* lies in the convex closure of the set $\{f(x,t)f'(x,t) : t\epsilon T_x\}$.

<u>Proof</u>. Suppose that 0 is not in the convex closure of the cited set. By the separation theorem there exists $z\epsilon E^{**} = E$ such that $f(x,t)\langle z,f'(x,t)\rangle > 0$ for all $t\epsilon T_x$. With no loss of generality we may

assume that $\|z\| = 1$. Since T_x is compact, there is an $\epsilon > 0$ less than all the numbers $f(x,t)\langle z, f'(x,t)\rangle$ for $t \epsilon T_x$. By the continuity properties of f and f' we may find a $\delta > 0$ such that for all $\|\zeta\| < \delta$, the following two conditions are satisfied:

$$\|f'(x + \zeta, t) - f'(x,t)\| < \tfrac{1}{2}\epsilon \qquad (t \epsilon T)$$

$$2f(x,t)\langle z, f'(x + \zeta, t)\rangle > \epsilon \qquad (t \epsilon T_x)$$

Now let $T_1 = \{t \epsilon T : |f(x,t)| > \frac{1}{2}\Delta(x)$ and $f(x,t)\langle z, f'(x,t)\rangle > \epsilon\}$. Clearly T_1 is an open set containing T_x. Thus $T_2 = T \sim T_1$ is a compact set containing no points of T_x. Consequently, the number $p = \sup_{t \epsilon T_2} |f(x,t)|$ is less than $\Delta(x)$. Select λ in accordance with the inequalities

(1) $\quad 0 < \lambda < \delta$

(2) $\quad \lambda Q < \Delta(x) - p \qquad$ where $Q = \frac{\epsilon}{2} + \sup_{t \epsilon T} |f'(x,t)|$

(3) $\quad \lambda Q^2 < \epsilon$.

We are going to prove that $\Delta(x - \lambda z) < \Delta(x)$. For $t \epsilon T_1$ we have $|f(x - \lambda z, t)|^2 = |f(x,t) - \lambda\langle z, f'(x - \Theta\lambda z, t)\rangle|^2 = |f(x,t)|^2 - 2\lambda f(x,t)\langle z, f'(x - \Theta\lambda z, t)\rangle + \lambda^2|\langle z, f'(x - \Theta\lambda z, t)\rangle|^2 < \Delta^2(x) - 2\lambda\epsilon + \lambda^2 Q^2 < \Delta^2(x) + \lambda[-2\epsilon + \epsilon]$. On the other hand, for $t \epsilon T_2$ we have $|f(x - \lambda z, t)| \le |f(x,t)| + \lambda|\langle z, f'(x - \Theta\lambda z, t)\rangle| \le p + \lambda Q < p + \Delta(x) - p = \Delta(x)$. ∎

<u>Remark</u>. For this theorem the function f may be defined on $S \times T$ (where S is an open subset of E) instead of on $E \times T$.

References

1. N. Dunford and J. Schwartz, Linear Operators, Part I, Interscience Publishers, New York, 1958.

2. N. Bourbaki, Espaces vectoriels topologiques, chapters 1 to 5, Hermann and Cie, Paris, 1953.

3. J. L. Kelley, General Topology, D. van Nostrand, New York, 1955.